

Tilburg University

Regulating Global Externalities

Heijmans, Roweno J.R.K.; Gerlagh, Reyer

Publication date:
2019

Document Version
Early version, also known as pre-print

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):
Heijmans, R. J. R. K., & Gerlagh, R. (2019). *Regulating Global Externalities*. (CentER Discussion Paper; Vol. 2019-001). CentER, Center for Economic Research.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

No. 2019-001

REGULATING GLOBAL EXTERNALITIES

By

Roweno J.R.K. Heijmans, Reyer Gerlagh

10 January 2019

ISSN 0924-7815
ISSN 2213-9532

Regulating Global Externalities

Roweno J.R.K. Heijmans and Reyer Gerlagh*

January 7, 2019

The most recent version of our paper can be found here:

www.dropbox.com/s/lryz9876oxewn86/Heijmans_Gerlagh_2019.pdf?dl=0

Abstract

The question in which we are interested is how a market inhabited by multiple agents, about whom we are differentially uncertain, and who trade goods the use of which imposes a negative effect on others, is to be ideally regulated. We show that a priori asymmetric uncertainty, when combined with a posteriori observed outcomes, is a rich source of information that can be used to reduce aggregate uncertainty. The observation implies that whereas asymmetric information usually entails a cost on welfare, it can help achieve greater efficiency in regulation.

JEL codes: D82, D83, H23

Keywords: asymmetric information, regulatory instruments, policy updating, asymmetric uncertainty, decision making under uncertainty

1 Introduction

One evening in a small bar, a male and a female sit by a table, smoking one cigarette after another. At the end of a long evening, the bar is black with smoke. As the couple leaves the establishment, some remaining customers voice their frustrations to the bartender. “We did not come here,” they say, “to drink our beers in the smoke of others. Please, for the sake of us all and the common good, next time do intervene.” These complainants being regulars, the bartender takes their discontent seriously. Counting the butts in the ashtray,

*Heijmans (r.j.r.k.heijmans@uvt.nl) and Gerlagh (r.gerlagh@uvt.nl) are at the Economics Department of Tilburg University.

he concludes the couple smoked 20 cigarettes that evening. As it happens, the female has frequented this bar for years. The bartender knows she always smokes 10 cigarettes on a night out. This leads to the conclusion that the male, who is new in town, has smoked 10 cigarettes as well. After elaborate calculations – into which we shall not dwell any further here – on the back of a coaster, the tapster decides an outcome where half this number of cigarettes, i.e. 10, is smoked is, in expectations, economically efficient for the full clientèle.

Next Tuesday, the same people visit the bar, and the barkeeper politely asks the couple to set a ceiling to their habit. “You can smoke five cigarettes each,” they are told, “though you can freely trade these right between the two of you, as long as the total does not exceed ten.” That evening, the bartender notices something quite peculiar. As the couple spends a cozy evening, the female smokes nine cigarettes, while the male consumes only one. What should the bartender do? He concludes the male must somehow have less of an appetite for smoking than initially thought. Following basic economic reasoning and in search of an outcome that is in the best overall interest, one week later the couple is asked to decrease their common budget further from ten to six cigarettes.

At a certain level of conceptual abstraction, the barkeeper’s situation is similar to a signal extraction problem, which as a branch of statistical theory goes by the name of filtering. Practitioners of the economic profession have applied these methods – and most frequently the Kalman filter (which we know to be the optimal filter when signals are normally distributed) – in many corners of their discipline; from the navigation of the Apollo project, to investment decisions (Townsend, 1983) and macroeconomic theory (Lucas, 1978), finance (Makarov and Rytchkov, 2012), time series econometrics (Hamilton, 1989; Baxter and King, 1999; Harvey and Trimbur, 2003; Talmon and Coifman, 2013) and behavioral economics (Mullainathan, 2002; Moore and Healy, 2008). Our mission will be to apply the insights so fruitfully used in all said applications and further develop these for the field of regulation.

The classical argument to favor markets is of remarkably individualistic nature. Any distribution of production or consumption that comes about through free exchanges in well-functioning market environments, so it goes, will always be at least as efficient as had allocations been set to individual agents directly. This is why many economists favor market-wide constraints over individual-specific ones. We do not doubt the rightness of this conviction, yet we do point out it is in a sense incomplete. Though markets help the efficiency of the *individual* distribution of some undertaken activity, they do not resolve the problem of joint efficiency at the aggregate.

Our central thesis builds on the logic presented above: when left to free exchanges the

agents in a market will trade and barter until no further mutually beneficial deals can be made, ensuring an efficient allocation of, say, cigarettes across individual agents. However, there is more to be inferred – about the aggregate efficiency – and the logic is simple. Suppose we observe some final distribution of cigarettes comes about and suppose also that this distribution is different from the most likely outcome we had initially assumed would occur. We realize that our initial estimation of smokers’ preferences need some correction. This is nothing new, some would argue, for this possibility is the very reason a market was established in the first place. However, rather than applaud ourselves for having achieved greater welfare by establishing a market, we admit that these wrong initial estimates may fairly well imply the aggregate number of cigarettes allowed to the market was, in fact, wrong, and a better estimate can be constructed on the basis of market evidence.

Consider the example with which we started. Suppose we know the preferences of smoker 1 while those of smoker 2 are uncertain but we initially expect them to be the same as for smoker 1. We allocate a total of ten cigarettes to the market, which is optimal given our initial estimate of preferences. This estimate also tells us that in fact each smoker will consume five cigarettes. Upon seeing that in reality smoker 1 smokes nine cigarettes while smoker 2 consumes only one, we realized our estimate of smoker 2’s preferences was wrong, because he or she is in fact a less devout consumer of the tobacco. Given we know the preferences of smoker 1, and given we now learned those of smoker 2 are stronger, we conclude ‘aggregate’ preferences for smoking are lower than expected. Standard economics then dictates that a lower aggregate cigarette consumption would consequently be socially optimal. Instead of 10, the bar tender gives them 6 cigarettes to smoke jointly.

In practical terms we are saying that the market cap on some activity with an associated externality should be endogenous to market behavior as the latter represents information regarding relevant characteristics of the market we are regulating and about which we are uncertain. If our exercise is to not only make the cap endogenous but in fact make it optimally so, we introduce a technical challenge. The standard problem, meaning with a fixed cap, imposes total consumption of cigarettes as exogenously given and lets the market sort out the individual allocations, subject to a simple constraint: total smoking is not to exceed the cap. With an endogenous cap on the other hand, individual smokers still trade freely, but these trades have an effect on how much can be freely traded in the first place, meaning the cap is a function of trades. Our question is simple but fundamental: what would be the properties of a more efficient cap? Turning the question around, we can ask: which of the properties of trading with an exogenous cap are (clearly) suboptimal. One prominent candidate is the use of information. An optimal exogenously fixed cap is

efficient in expectations, before any observation on individual outcomes is available.

We propose to make use of more refined information, namely not aggregate but individual behavior. This information is readily available, as is the information on aggregate behavior, yet it provides a rich source for information about aggregate preferences that can greatly improve efficiency of regulation. The intuition we presented earlier in this introduction. What we want to emphasize is that we particularly exploit asymmetries in uncertainty. If the bar tender had exactly the same prior information about the female and male, he could update the estimates for individual preferences, but *not* for aggregate preferences. The aggregate revision requires asymmetry in priors.

Our approach has a flavor of mechanism design, but we regulate a market only in the aggregate, and thus we search for information about aggregate preferences. The approach followed by mechanism designers is to offer menus or contracts to each individual agent, searching for information on individual preferences with no mechanism for the aggregate regulation level.

A careful reading of the smoking example reveals another property of our regulatory policy: we focus on quantities only. It is well known at least since Roberts and Spence (1976) that a combination of instruments, such as a hybrid between prices and quantities, can in principle do better than one of these instruments alone. Say, the bar tender could have offered extra quota for some price. Again, we abstract away from such combinations of different instruments and instead focus on pure instruments and take this type of regulation as far as it can get. What we will see is that, when used smartly, welfare losses can be substantially mitigated compared to pure trading policies, motivating the idea that no need exists for complicated hybrid instruments if one is willing to carefully contemplate on and improve the pure instruments.

2 Model

Our universe is a bar, inhabited by two atomic or price-taking smokers and a handful of other clients.¹ To every smoker i , smoking \tilde{s}_i cigarettes yields private benefits $B_i(\tilde{s}_i; \theta_i)$. The parameter θ_i can be thought of as a preference shifter, affecting how much pleasure is derived from smoking a given number of cigarettes and is known to smoker i but not to any other individual. Although the actual realization of θ_i is private information, it is

¹A price-taking duo may sound a wee bit unrealistic to many economists. One possible interpretation is that our environment is not a bar, but instead a large-scale disco in New York City where smoking is allowed but regulated, and where we assume there are two large groups of smokers. In this case, every individual smoker only has a negligible effect on the eventual price of cigarettes.

common knowledge that $\mathbb{E}[\theta_i] = 0$, $\mathbb{E}[\theta_i^2] = \sigma_i^2$, and $\mathbb{E}[\theta_1\theta_2] = \rho\sigma_1\sigma_2$. Because this universe is uncertain to the extent that θ_i cannot be properly predicted, σ_i^2 is a logical measure for the uncertainty about a agent i 's preferences. To say that uncertainty is asymmetric in our terminology is equivalent to saying that $\sigma_1 \neq \sigma_2$.

As far as smoking emits smoke, it is disliked by the other customers. We will assume that the severity of this externality depends on the total amount of smoke in the bar only, independent of which smoker exhaled it, so we are dealing with a global externality. Because there is (in this simplified universe) a one-to-one relationship between the amount of cigarettes consumed and the amount of smoke produced, we can treat the externality as a cost, broadly interpreted, $C(\tilde{s}_1 + \tilde{s}_2)$ depending solely on the aggregate number of cigarettes smoked.

As the other clients are bothered by the amount of smoke emitted, the barkeeper faces the task of finding quantities \tilde{s}_1 and \tilde{s}_2 that maximize social welfare:

$$W = B_1(\tilde{s}_1; \theta_1) + B_2(\tilde{s}_2; \theta_2) - C(\tilde{s}_1 + \tilde{s}_2). \quad (1)$$

In the absence of asymmetric information, the fully knowledgeable barkeeper can set these quantities directly or else put a price on cigarettes that will make the individual smokers consume the same quantities, and these two instruments are perfectly equivalent, see Montgomery (1972). As was first shown by Weitzman (1974), this formal equivalence between instruments breaks down once we introduce an informational disparity, captured here by θ_i .

It will serve the analysis to make some restrictive assumptions regarding the forms benefits and costs take. In particular, we shall consider both to be of the linear-quadratic form. Thus, benefits to smoker i are given by:

$$B_i(\tilde{s}_i; \theta_i) = (p_i^* + \theta_i)(\tilde{s}_i - s_i^*) - \frac{\beta_i}{2}(\tilde{s}_i - s_i^*)^2, \quad (2)$$

where the vector (p_i^*, s_i^*) is common knowledge and will be elaborated upon soon. Marginal benefits are therefore linear in cigarettes with the intercept determined by θ_i :

$$MB_i(\tilde{s}_i) = p_i^* - \beta_i(\tilde{s}_i - s_i^*) + \theta_i. \quad (3)$$

Costs as a result of disutility from smoke are described by the functional form:

$$C(\tilde{s}_1 + \tilde{s}_2) = p^*(\tilde{s}_1 + \tilde{s}_2 - s_1^* - s_2^*) + \frac{\gamma}{2}(\tilde{s}_1 + \tilde{s}_2 - s_1^* - s_2^*)^2. \quad (4)$$

Marginal costs due to smoke are then seen to be:

$$MC(\tilde{s}_1 + \tilde{s}_2) = p^* + \gamma(\tilde{s}_1 + \tilde{s}_2 - s_1^* - s_2^*). \quad (5)$$

For brevity of notation, where convenient we may write $\tilde{S} = \tilde{s}_1 + \tilde{s}_2$ and $S^* = s_1^* + s_2^*$. Our model now has eight parameters $(\beta_i, \gamma, p_i^*, s_i^*, p^*)$ describing the slopes and intercepts of three linear curves. We need only two per curve, that is six in total. Consequently, we may take the freedom to reduce the number of parameters by defining $p^* = p_1^* = p_2^*$, with the convenient implication that (p^*, s_1^*, s_2^*) is the vector of welfare-maximizing prices and cigarettes for smoker i given preferences turn out as expected ($\theta_1 = \theta_2 = 0$). We label this the ex-ante optimum. It is easily seen that global marginal costs and regional marginal benefits equal p^* . This is clearly not an assumption, nor even a normalization; it is a definition. Entailing no more than a simplification of otherwise potentially cumbersome notation, we introduce it here for our own convenience without any implicated loss of generality.

Before proceeding to the analysis, we introduce some further notation. Superscripts will be scenario labels for equilibrium outcomes. Moreover, let \tilde{x}^k denote the value of a variable x under policy k , then let $x^k := \tilde{x}^k - x^*$ be the deviation of x under policy k from the ex-ante expected optimal value x^* , and let $\Delta^k x := \tilde{x}^k - x^{SO}$ denote the difference between the value of x under scenario k and its ex post socially optimal value (to be derived shortly).

The game has the following stages:

1. The barkeeper chooses an instrument to regulate the market for cigarettes.
2. Both smokers observe their individual preference shock θ_1 and θ_2 .
3. Trade clears the market. Prices and quantities are chosen, jointly for both smokers, consistent with utility maximization by each smoker,

$$-\beta_i s_i^k + \theta_i = p_i^k, \quad (6)$$

while the policy rules determine the relation between quantities and prices within and across smokers.

2.1 Symmetric Information and Perfect Foresight

By standard arguments, it is immediately clear that marginal benefits of smoking should equal the marginal costs of smoke in an efficient outcome which implies $MB_1 = MB_2$. Since marginal benefits also equal prices, these are the same, so $p_1^{SO} = p_2^{SO} = p^{SO}$. Labeling the symmetric information equilibrium as Social Optimum, we have the profit-maximization condition (6) and

$$\gamma(s_1^{SO} + s_2^{SO}) = p^{SO},$$

so that the Social Optimum is easily characterized:

$$p^{SO} = \frac{\gamma(\beta_2\theta_1 + \beta_1\theta_2)}{\gamma\beta_1 + \gamma\beta_2 + \beta_1\beta_2}, \quad (7)$$

$$s_i^{SO} = \frac{\beta_{-i}\theta_i + \gamma(\theta_i - \theta_{-i})}{\gamma\beta_1 + \gamma\beta_2 + \beta_1\beta_2}, \quad (8)$$

$$S^{SO} = \frac{\beta_2\theta_1 + \beta_1\theta_2}{\gamma\beta_1 + \gamma\beta_2 + \beta_1\beta_2}, \quad (9)$$

where $i \in \{1, 2\}$ and $-i$ is the complement. Thus, a positive preference shift induces increased consumption of cigarettes by the smoker to whom it occurred, and decreases it for the other, though aggregate consumption and the common price always increase for a positive shift to either smoker.

2.2 Implementing the Social Optimum: First-best Equilibrium

A straightforward mechanism that implements the Social Optimum is a simple ascending clock auction. In its most basic form, the barkeeper offers a supply curve, in price-cigarette space, that coincides with the marginal cost curve. This way, utility-maximizing smokers necessarily incorporate the externality caused by their smoke into smoking behavior, guaranteeing implementation of the Social Optimum as a first-best equilibrium.

That such an easy way of implementing the Social Optimum exists is reason for optimism. After all, it implies that the Social Optimum can be reached without great effort or complication. Nonetheless, practice suggests that such instruments as the ascending clock auction are not feasible for reasons outside the scope of our model. The barkeeper is in such a case constrained to other instruments, accepting a second-best regulated equilibrium as at least superior to no regulation at all. It is that context that we are about to define and analyze. Before doing so, however, we will first derive in general terms

the aggregate welfare loss under a given policy relative to the first-best, Social Optimum welfare level.

2.3 Welfare Costs of Policies

By definition of the difference under policy k with the social optimum and considering a smoker's equilibrium behavior (6), it is immediate that smoking deviations from the Social Optimum scale with price deviations:

$$\Delta^k p_i = -\beta_i \Delta^k s_i. \quad (10)$$

Welfare losses are then given by:

$$\begin{aligned} \Delta^k W &= \mathbb{E} [\Delta^k B_1 + \Delta^k B_2 - \Delta^k C_1 - \Delta^k C_2] \\ &= \frac{\gamma}{2} \mathbb{E} [(\Delta^k S)^2] + \sum_i \frac{\beta_i}{2} \mathbb{E} [(\Delta^k s_i)^2]. \end{aligned} \quad (11)$$

Given realized preferences, individual prices and cigarette consumption map injectively. Policies featuring equal prices across smokers thus admit the property that individual and aggregate consumption scale with the common price. Consequently, for such policies welfare losses can be written as a function of the price gap:

$$\Delta^k W = \frac{1}{2} \frac{(\gamma\beta_1 + \gamma\beta_2 + \beta_1\beta_2)(\beta_1 + \beta_2)}{\beta_1^2\beta_2^2} \mathbb{E} [(\Delta^k p)^2]. \quad (12)$$

3 Policies

3.1 Quotas

The simplest possible policy simply sets quotas for each smoker individually:

Definition 1 (Quotas). To both smokers individually, the barkeeper allocates the ex-ante optimal amount of cigarettes s_i^* :

$$s_1 = s_2 = 0, \quad (13)$$

while prices adjust to reach equilibrium on the market for cigarettes (6).

We readily obtain expected welfare losses:

$$\Delta^Q W = \frac{1}{2} \frac{(\gamma + \beta_2)\sigma_1^2 + (\gamma + \beta_1)\sigma_2^2 - 2\gamma\rho\sigma_1\sigma_2}{\gamma\beta_1 + \gamma\beta_2 + \beta_1\beta_2}. \quad (14)$$

For future reference, it is important to note that Quantities under autarky as regulation can be considered the execution of the welfare program

$$\max_{s_1, s_2} \mathbb{E} \quad W(s_1, s_2; \theta_1, \theta_2) \quad (15)$$

That is, Quantities is the optimal choice under the information constraint that both quantities must be set before any information is revealed, and without the use of any information extracted from markets. It admits the desirable property that expected marginal benefits equals marginal costs:

$$\mathbb{E}[MB_i | s_i] = MC \quad (16)$$

where we note that the RHS marginal costs are perfectly known, when quantities are set whereas the prices at the LHS are stochastic variables due unknown preferences θ_i .

3.2 Taxes

Another possibility is to impose a price or tax on cigarettes:

Definition 2 (Taxes). The regulator cigarette-taxes at the ex-ante optimal level $\tilde{p}_1 = \tilde{p}_2 = p^*$:

$$p_1 = p_2 = 0. \quad (17)$$

Cigarette consumption adjust to reach equilibrium on market for cigarettes (6).

Smokers consume cigarettes conditional on preference θ_i according to $\Delta^P s_i = \theta_i / \beta_i$. We now invoke (10) and obtain welfare loss:

$$\Delta^T W = \frac{1}{2} \frac{\gamma^2(\beta_1 + \beta_2)}{\beta_1^2 \beta_2^2} \frac{\beta_2^2 \sigma_1^2 + \beta_1^2 \sigma_2^2 + 2\beta_1 \beta_2 \rho \sigma_1 \sigma_2}{\gamma\beta_1 + \gamma\beta_2 + \beta_1\beta_2}. \quad (18)$$

Taxes as a policy can be considered the execution of the welfare program

$$\max_{p_1, p_2} \mathbb{E} \quad [W(s_1, s_2; \theta_1, \theta_2)] \quad (19)$$

$$\text{s.t.} \quad -\beta_i s_i^k + \theta_i = p_i^k \quad (6)$$

That is, Taxes are the optimal choice under the informational constraint that cigarette-prices for both smokers must be set before any information is revealed. It has the desirable property that marginal benefits equals expected marginal costs:

$$MB_1 = MB_2 = \mathbb{E}[MC|p] \quad (20)$$

where we note that the LHS are marginal benefits which are known with certainty when Taxes are set.

4 The Market For Cigarettes

The barkeeper may well understand that through establishing a market for cigarettes even better outcomes, in terms of social welfare, can be obtained. The basic idea is very simple and familiar to every economist in the modern tradition: when a market for cigarettes exists, smokers can freely exchange their smokes, and since such exchanges will only take place as long as they are mutually beneficial, aggregate welfare will be larger as compared to a scenario where no market exists. The policy where a market with the essential feature of free exchange is created will be called Trading.

4.1 Trade

Definition 3 (Trading). The barkeeper allocates the ex-ante optimal number of cigarettes s_i^* to both smokers. smokers can freely exchange cigarettes, subject to:

$$s_1 + s_2 = 0. \quad (21)$$

Equilibrium on the market for cigarettes implies (6). Optimization and free trading ensures cigarettes are exchanged until marginal benefits are equal for both smokers:

$$p_1 = p_2. \quad (22)$$

Trading allows smokers the flexibility to efficiently redistribute cigarette allocations in response to shocks, subject to the constraint that total smoking is fixed. It follows:

$$\Delta^T W = \frac{1}{2} \frac{1}{\beta_1 + \beta_2} \frac{\beta_2^2 \sigma_1^2 + \beta_1^2 \sigma_2^2 + 2\beta_1 \beta_2 \rho \sigma_1 \sigma_2}{\gamma \beta_1 + \gamma \beta_2 + \beta_1 \beta_2}. \quad (23)$$

The following proposition is now immediate:

Proposition 1. *Trading always outperforms Quotas in terms of welfare.*

Proof. In Appendix. Q.E.D.

It is quite trivial that Trading improves welfare compared to autarkic Quotas. Under the former, aggregate smoking, and therefore costs, are fixed. Yet through a process of mutually beneficial exchanges, benefits including revenues from cigarette sales, increase for both smokers, raising welfare overall.

The conceptual quality of trade as regulatory principle can be seen more clearly when taking a more principal look at the policy and its rules. Quotas under autarky as policy can be considered the implementation of welfare program (15). Trading, on the other hand, is the execution of:

$$\max_G \mathbb{E} \left[\max_{s_1, s_2} W(s_1, s_2; \theta_1, \theta_2) \right] \quad (24)$$

$$\text{s.t. } s_1 + s_2 = G \quad (25)$$

That is, Trading effectively delays the choice for optimal smoking per smoker until the point information is revealed. It admits the desirable property that not only do marginal costs equal expected marginal benefits (which is true also for Quotas, see (16)) – realized marginal benefits are also equal for both smokers:

$$MB = MB_1 = MB_2 \quad (26)$$

$$\mathbb{E}[MB|s_1 + s_2] = MC \quad (27)$$

One might wonder about price volatility in the market for cigarettes. After all, if welfare under Trading is higher than under Quotas, as Proposition 1 establishes, and if we know from condition (22) that prices will be equal for both smokers, given also it is known that in markets with constant prices welfare losses scale with price deviations (equation (12)), then should it not follow price volatility is lower under Trading than under Quotas? The proposition maintained a purposeful silence about aggregate price volatility in the Trading market. It did so because results are mixed. One might at first suspect trade to reduce global volatility compared to autarkic Quotas, but a simple thought experiment is illustrative for the wrongness of this intuition. For that, we note first that prices in equilibrium equate marginal benefits, so that the volatility of prices is also equal, in equilibrium, to the volatility of marginal benefits. Consider then two smokers, the second with more uncertain preferences than the first, $\sigma_2 \gg \sigma_1$, as well as a larger absorptive capacity, $\beta_2 \gg \beta_1$. Consequently, if under Quotas, say, five cigarettes

are allocated to both of them, we expect price volatility to be almost zero for smoker 1 but much larger for smoker 2, for only the latter experiences large ‘preference shifts’. When we introduce Trading, the globally efficient allocation of cigarettes may be such that many flow to smoker 2, inducing a much stronger price fluctuation for smoker 1 than under Quotas. This means that the first smoker has practically *imported* part of the price volatility from the second. As the first is characterized by little absorption, prices in equilibrium under Trading will mainly be driven by the second smoker. Aggregate price volatility has in fact increased.

This point is neither just a footnote nor a mere theoretical curiosity. For the success of a Trading regime, it is of fundamental importance. After all, would a stable smoker be willing to expose itself directly to the risk of its unstable trading partners? It appears, to say the least, unlikely. There is no trivial solution to this fundamental problem. Yet we are about to show how it can be substantially mitigated.

4.2 Rates of substitution

True as it may be that Trading is good for welfare, we argue that it is not best, and therefore the barkeeper can do better. In fact, we will shortly show that the regulator is able to do so rather easily, although some delicate insights must be developed first, for else one may not fully understand the subtle mechanisms at work.

The primary notion we need to establish is that exist not a single marginal rate of substitution for permits; rather, there are two. One operates at the individual smoker’s level, the other on the aggregate or global level.

The individual rate, labeled MRS_i , appears to be the rate one most frequently has in mind when speaking loosely of marginal substitution, it being the rate at which cigarettes can change hands between individual smokers. Although it might seem intuitive that manipulating this trading ratio’ could improve welfare, it can in fact be shown that letting it deviate from unity is never optimal. This follows from first principles. An efficient allocation of cigarettes equates marginal benefits across smokers. Smokers, however, only have an incentive to perfectly equate marginal benefits (or the value of smoking) when cigarettes can be traded one-to-one, that is, when one cigarette lost by a smoker translates into exactly one cigarette gained by the other. Therefore, only if the MRS_i is unity will smokers have no incentive to trade cigarettes other than the full equalization of marginal benefits.

These observations do not imply, as might come across at first, that the barkeeper’s arsenal of instruments is left depleted. Individual trades must be left untouched, it is

true, but *aggregate* trades can still be manipulated. Indeed, one cannot seriously think of any proper economic reasoning contra such operations. It combines the best of two worlds: individual smokers consider their impact on aggregate trade flows to be negligible and will therefore trade freely on the basis of one-to-one exchanges, steering cigarette consumption toward a perfect equalization of marginal benefits, while at the global level smoking is nonetheless adapted to correct for the possibly flawed and suboptimal initial endowment of cigarettes revealed by the market through its self-chosen trade flows. In its simplest form, this manipulation of aggregate permit trades operates through a fixed ratio, called the aggregate marginal rate of substitution, MRS_a for short. To find a handle on the aggregate rate of substitution, we will formulate regulation as the solution to a welfare-maximization problem. We then replicate Trading as regulation, and find a natural generalization, which we label Stabilized Trading. That approach will inform us how to model such aggregate manipulations, deriving how these should occur in an optimal bar.

4.3 Stabilized Trading

Definition 4 (Stabilized Trading). The barkeeper adapts the aggregate allocation of cigarettes based on observed trade for fixed $MRS_A = \delta$:

$$\delta s_1 + s_2 = 0. \quad (28)$$

Profit maximization and free trading with $MRS_i = 1$ ensures that smokers allocate cigarettes so that marginal productivity is equal for both of them:

$$p_2 = p_1. \quad (29)$$

We observe that for $\delta = 1$, Stabilized Trading is one-to-one also at the aggregate level and thus equivalent to traditional Trading. This observation immediately suggests that Stabilized Trading always outperforms traditional Trading, since the barkeeper is free to choose a stabilization rate δ equal to unity but not imposed to do so, which added freedom cannot deteriorate global welfare. We can derive the following result:

Proposition 2. *The optimal stabilization rate is given by:*

$$\delta^* = \frac{\beta_1[\sigma_2^2 - \rho\sigma_1\sigma_2] + \gamma[\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2]}{\beta_2[\sigma_1^2 - \rho\sigma_1\sigma_2] + \gamma[\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2]}, \quad (30)$$

Stabilized Trade equaling Trade is a measure-zero event. In particular,

$$\delta^* \begin{matrix} \leq \\ > \end{matrix} 1 \iff \frac{\beta_1}{\beta_2} \begin{matrix} \leq \\ > \end{matrix} \frac{\sigma_1^2 - \rho\sigma_1\sigma_2}{\sigma_2^2 - \rho\sigma_1\sigma_2}. \quad (31)$$

Before we turn our attention to the particular properties of the stabilization rate, we want to appreciate its principles more fundamentally at the informational level. If we generally consider instruments from the class of Quantity-based ones, we understand these can be interpreted as allowing, in principle, full observation of each smoker's cigarette consumption and setting restrictions to these. Thus, what we aim for as the most efficient quantity-based instrument is an equalization of marginal benefits and marginal costs, after observing all quantities (cf (27))!

$$\mathbb{E}[MB|s_1, s_2] = MC \quad (32)$$

Note that upon observing both quantities, the difference in preference shifts can be constructed:

$$\mu \equiv \theta_2 - \theta_1 = \beta_1 s_1 - \beta_2 s_2 \quad (33)$$

Using the demand equation and plugging in μ , we find

$$\mathbb{E}[MB|s_1, s_2] = \mathbb{E}[\theta_1|\mu] - \beta_1 s_1 \quad (34)$$

$$= \mu \frac{\mathbb{E}[\mu\theta_1]}{\mathbb{E}[\mu^2]} - \beta_1 s_1 \quad (35)$$

$$= \mu \frac{\rho\sigma_1\sigma_2 - \sigma_1^2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2} - \beta_1 s_1 \quad (36)$$

$$= -\frac{\sigma_2^2\rho - \sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}\beta_1 s_1 - \frac{\sigma_1^2 - \rho\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}\beta_2 s_2 \quad (37)$$

The RHS is straightforwardly equated to marginal damages:

$$\gamma(s_1 + s_2) = -\frac{\sigma_2^2\rho - \sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}\beta_1 s_1 - \frac{\sigma_1^2 - \rho\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}\beta_2 s_2, \quad (38)$$

which for convenience we rewrite as

$$\delta s_1 + s_2 = 0, \quad (28)$$

with δ given by

$$\delta = \frac{\gamma + \frac{\sigma_2^2 \rho - \sigma_1 \sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho \sigma_1 \sigma_2} \beta_1}{\gamma + \frac{\sigma_1^2 - \rho \sigma_1 \sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho \sigma_1 \sigma_2} \beta_2} \quad (39)$$

$$= \frac{\gamma(\sigma_1^2 + \sigma_2^2 - 2\rho \sigma_1 \sigma_2) + (\sigma_2^2 - \rho \sigma_1 \sigma_2) \beta_1}{\gamma(\sigma_1^2 + \sigma_2^2 - 2\rho \sigma_1 \sigma_2) + (\sigma_1^2 - \rho \sigma_1 \sigma_2) \beta_2}. \quad (40)$$

This fundamental approach toward constructing Stabilized Trading allows for an equally fundamental insight: Stabilized Trading and its optimal stabilization rate is not ‘just another’ regulation rule that is optimized, it is *the* most efficient quantity-based regulation.

Several interesting properties of the optimal stabilization rate deserve elaboration. First, as stated in the Proposition, hardly ever will this rate be unity. Thus, Stabilized Trading outperforms traditional Trading in terms of welfare almost always. Second, the optimal stabilization rate may well be negative, meaning that higher-than-expected smoking of one smoker translates into higher-than-expected smoking by the other smoker too. Closer inspection reveals such is more likely to occur for strongly positively correlated preferences and very asymmetric uncertainty. This is easily understood: if we do not know the second smoker too well but we do know he is very much alike the first smoker, then smoking should be increased for both, or for none. A negative stabilization rate bears some resemblance with putting negative weights on observations in making (econometric) predictions (see Bunn, 1985; Elliott and Timmermann, 2004; Timmermann, 2006).

Third, the share of preference shifts absorbed by a smokers decreases in the smoker’s responsiveness of benefits to smoking, that is, in its slope parameter β . For any adaption of global smoke to shocks, marginal costs change accordingly. Since trade leads to the ex-post equality of individual marginal benefits, and since an optimal mechanism equates individual marginal benefits to global marginal costs, for any realized pair of shifts smoking change relatively less for the smoker with steeper marginal benefits.

4.4 Asymmetric Uncertainty

Finally, our potentially most interesting observation is that the stabilization rate tends to increase, all else equal, if the uncertainty of smoker 2 increases. This is clearly driven by our choice of definition; had we alternatively set $s_1 + \delta s_2 = 0$, the result would be reversed. The intuition remains unaffected, though, being that most variability is warranted for the smoker who can be learned about most.

This intuition is most clearly understood by considering a rather extreme example.

Suppose that smoker 1 is so frequent a visitor of this bar that the barkeeper in fact has $\sigma_1 = 0$, i.e. smoker 1's preferences are perfectly predictable without error. Assume moreover that this is not true for smoker 2, the new smoker in town, meaning $\sigma_2 > 0$. The barkeeper thus clearly faces a situation with asymmetric uncertainty about the smokers' preferences.

Suppose now the barkeeper observes a reallocation of the initial distribution of cigarettes, such that smoker 2 now smokes only four and smoker 1 a total of six. Since the barkeeper knows the preferences of smoker 1 perfectly, and since trade will lead to equality of marginal benefits, the barkeeper can thus infer the precise marginal benefits also of smoker 2. But this, in turn, implies the barkeeper is able to identify the exact preferences of smoker 2, including, that is, the initially unobserved shift θ_2 .

We learn something important here. Without paying attention to an asymmetry of uncertainty, as in the classical argument made in favor of trade and markets, upon seeing a flow of cigarettes from smoker 2 to smoker 1 all the barkeeper could conclude is that preferences had shifted in such a way that the value of an extra cigarette was (much) higher for smoker 1 than for smoker 2. The barkeeper may feel happy, for has not trade resulted in an efficient exchange of cigarettes? True, these trades are efficient, but only at the *individual* smoker's level. The new allocation of cigarettes is constrained Pareto optimal, i.e. given a potentially suboptimal aggregate consumption of cigarettes the individual consumption levels are optimal. The barkeeper could have stopped there. However, we argue it could have been known that the aggregate amount of smoking was in fact inefficient. Taking into account the fact that smoker 1 has far less uncertain preferences than smoker 2, upon seeing a flow of cigarettes from the latter to the former, it can be concluded that the 'aggregate' or 'average' preference for smoking is in fact lower than anticipated. After all, a sale of cigarettes from smoker 2 to smoker 1 tells us more than simply that our initial expectations about individual preferences were off – it tells us that smoker 2 has a weaker preference for cigarettes than smoker 1, and since the latter's preferences are known with certainty, we can conclude what we may loosely write as $\theta_2 < \mathbb{E}[\theta_2]$. The economist then knows enough to conclude that the aggregate cap on smoking shall be tightened.

The lesson drawn from this extreme example is of course more generally true and applicable: there is *more* scope for learning about agent whose preferences are *more* uncertain. Etc.

Given an optimal stabilization rate, we can solve for the associated level of expected global welfare. We do so in the our Theorem:

Theorem 1. *Stabilized Trading is strictly welfare-superior to Quotas and Trading, with*

welfare given by:

$$\Delta^{ST}W = \frac{1}{2} \frac{\beta_1 + \beta_2}{\gamma\beta_1 + \gamma\beta_2 + \beta_1\beta_2} \frac{(1 - \rho^2)\sigma_1^2\sigma_2^2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}. \quad (41)$$

Proof. In Appendix. Q.E.D.

Since $\Delta^TW > \Delta^{ST}W$, we refer to (12) and conclude that Stabilized Trading moves prices closer to the Social Optimum than traditional Trading. We also see the illustration used above reflected in our theorem. If the barkeeper has perfect information on one smoker, $\sigma_1 = 0$, then all preferences are revealed through trade and no welfare losses occur. Moreover, it is straightforward to derive that price volatility relative to the *ex ante* price level for both smokers is also lower under Stabilized Trading:

Proposition 3. *Stabilized Trading admits lower price volatility than Trading:*

$$\forall i : \mathbb{E} \left[(p^{ST})^2 \right] \leq \mathbb{E} \left[(p^T)^2 \right]. \quad (42)$$

Proof. In appendix. Q.E.D.

5 Discussion and Conclusions

It has herein been demonstrated that fairly simple and straightforward manipulations of traditional free trade in a market under regulation can yield substantial welfare gains. A key concept used thereto is that of asymmetric uncertainty and the implied differential learning potentials about different regulated agents. Our results can be directly applied to all activities where a global externality creates market imperfections. One example would be emission trading systems or regions, such as the different sectors or countries jointly regulated in EU ETS or US states in RGGI.

References

- Baxter, M. and King, R. G. (1999). Measuring business cycles: approximate band-pass filters for economic time series. *Review of economics and statistics*, 81(4):575–593.
- Bunn, D. W. (1985). Statistical efficiency in the linear combination of forecasts. *International Journal of Forecasting*, 1(2):151–163.

- Elliott, G. and Timmermann, A. (2004). Optimal forecast combinations under general loss functions and forecast error distributions. *Journal of Econometrics*, 122(1):47–79.
- Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica: Journal of the Econometric Society*, pages 357–384.
- Harvey, A. C. and Trimbur, T. M. (2003). General model-based filters for extracting cycles and trends in economic time series. *Review of Economics and Statistics*, 85(2):244–255.
- Lucas, R. E. (1978). Asset prices in an exchange economy. *Econometrica: Journal of the Econometric Society*, pages 1429–1445.
- Makarov, I. and Rytchkov, O. (2012). Forecasting the forecasts of others: Implications for asset pricing. *Journal of Economic Theory*, 147(3):941–966.
- Montgomery, W. D. (1972). Markets in licenses and efficient pollution control programs. *Journal of Economic Theory*, 5(3):395–418.
- Moore, D. A. and Healy, P. J. (2008). The trouble with overconfidence. *Psychological review*, 115(2):502.
- Mullainathan, S. (2002). A memory-based model of bounded rationality. *The Quarterly Journal of Economics*, 117(3):735–774.
- Roberts, M. J. and Spence, M. (1976). Effluent charges and licenses under uncertainty. *Journal of Public Economics*, 5(3-4):193–208.
- Talmon, R. and Coifman, R. R. (2013). Empirical intrinsic geometry for nonlinear modeling and time series filtering. *Proceedings of the National Academy of Sciences*, 110(31):12535–12540.
- Timmermann, A. (2006). Forecast combinations. *Handbook of economic forecasting*, 1:135–196.
- Townsend, R. M. (1983). Forecasting the forecasts of others. *Journal of Political Economy*, 91(4):546–588.
- Weitzman, M. L. (1974). Prices vs. quantities. *Review of Economic Studies*, 41(4):477–491.

A Derivations and Proofs

DERIVATION OF (23):

Combining the definition with the firms' FOCs, (6), we find the change in permit use by region:

$$\Delta^T_{s_1} = \frac{\theta_1 - \theta_2}{\beta_1 + \beta_2} \quad (43)$$

$$\Delta^T_{s_2} = \frac{\theta_2 - \theta_1}{\beta_1 + \beta_2}. \quad (44)$$

PROOF OF PROPOSITION 1:

Proof. For the first part, note that Trading outperforms Quantities if and only if the following condition is satisfied:

$$\begin{aligned} \frac{\beta_2^2 \sigma_1^2 + \beta_1^2 \sigma_2^2 + 2\beta_1 \beta_2 \rho \sigma_1 \sigma_2}{\beta_1 + \beta_2} &< (\gamma + \beta_2) \sigma_1^2 + (\gamma + \beta_1) \sigma_2^2 - 2\gamma \rho \sigma_1 \sigma_2 \\ &\iff \\ \frac{2\rho \sigma_1 \sigma_2}{\sigma_1^2 + \sigma_2^2} &< \frac{\gamma \beta_1 + \gamma \beta_2 + \beta_1 \beta_2}{\gamma \beta_1 + \gamma \beta_2}, \end{aligned}$$

which is always true.

Q.E.D.

DERIVATION OF (30):

Regional and global deviations from Socially Optimal permit use are given by:

$$\Delta^{ST}_{s_1} = \frac{\beta_2}{\beta_1 + \delta \beta_2} \frac{[\delta \beta_2 - \gamma(1 - \delta)]\theta_1 + [\beta_1 + \gamma(1 - \delta)]\theta_2}{\gamma \beta_1 + \gamma \beta_2 + \beta_1 \beta_2} \quad (45)$$

$$\Delta^{ST}_{s_2} = \frac{\beta_1}{\beta_1 + \delta \beta_2} \frac{[\delta \beta_2 - \gamma(1 - \delta)]\theta_1 + [\beta_1 + \gamma(1 - \delta)]\theta_2}{\gamma \beta_1 + \gamma \beta_2 + \beta_1 \beta_2} \quad (46)$$

$$\Delta^{ST}_Q = \frac{\beta_1 + \beta_2}{\beta_1 + \delta \beta_2} \frac{[\delta \beta_2 - \gamma(1 - \delta)]\theta_1 + [\beta_1 + \gamma(1 - \delta)]\theta_2}{\gamma \beta_1 + \gamma \beta_2 + \beta_1 \beta_2}. \quad (47)$$

Define

$$\xi := \frac{\beta_1 + \gamma(1 - \delta)}{\beta_1 + \delta \beta_2} \implies 1 - \xi := \frac{\delta \beta_2 - \gamma(1 - \delta)}{\beta_1 + \delta \beta_2}. \quad (48)$$

Welfare losses can now be written as:

$$\begin{aligned}\Delta^{ST}W &= \frac{1}{2} \frac{\gamma(\beta_1 + \beta_2)^2 + \beta_1^2\beta_2 + \beta_1\beta_2^2}{(\gamma\beta_1 + \gamma\beta_2 + \beta_1\beta_2)^2} \mathbb{E}[(1 - \xi)\theta_1 + \xi\theta_2]^2 \\ &= \frac{\beta_1 + \beta_2}{2} \frac{(1 - \xi)^2\sigma_1^2 + \xi^2\sigma_2^2 + 2\xi(1 - \xi)\rho\sigma_1\sigma_2}{\gamma\beta_1 + \gamma\beta_2 + \beta_1\beta_2}.\end{aligned}\quad (49)$$

If for notational convenience, we define:

$$\psi := \frac{1}{2} \frac{\beta_1 + \beta_2}{\gamma\beta_1 + \gamma\beta_2 + \beta_1\beta_2}, \quad (50)$$

it is straightforward to derive:

$$\frac{\partial}{\partial \xi} \frac{\Delta^{ST}W}{\psi} = 2\xi\sigma_2^2 - 2(1 - \xi)\sigma_1^2 + 2(1 - \xi)\rho\sigma_1\sigma_2 - 2\xi\rho\sigma_1\sigma_2. \quad (51)$$

The welfare-maximizing ξ^* therefore satisfies:

$$\xi^* = \frac{\sigma_1^2 - \rho\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}. \quad (52)$$

From the definition of ξ , the optimal stabilization rate δ^* follows:

$$\delta^* = \frac{(\beta_1 + \gamma)[\sigma_2^2 - \rho\sigma_1\sigma_2] + \gamma[\sigma_1^2 - \rho\sigma_1\sigma_2]}{(\beta_2 + \gamma)[\sigma_1^2 - \rho\sigma_1\sigma_2] + \gamma[\sigma_2^2 - \rho\sigma_1\sigma_2]}, \quad (53)$$

as stated.

PROOF OF THEOREM 1:

Proof. Plugging (52) in (49), we find:

$$\begin{aligned}\frac{\Delta^{ST}W}{\psi} &= \left[\frac{\sigma_2^2 - \rho\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2} \right]^2 \sigma_1^2 + \left[\frac{\sigma_1^2 - \rho\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2} \right]^2 \sigma_2^2 \\ &\quad + \left[\frac{\sigma_2^2 - \rho\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2} \right] \left[\frac{\sigma_1^2 - \rho\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2} \right] \rho\sigma_1\sigma_2 \\ &= \frac{(1 - \rho^2)\sigma_1^2\sigma_2^2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2} \\ &\implies \\ \Delta^{ST}W &= \frac{1}{2} \frac{\beta_1 + \beta_2}{\gamma\beta_1 + \gamma\beta_2 + \beta_1\beta_2} \frac{(1 - \rho^2)\sigma_1^2\sigma_2^2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2},\end{aligned}$$

as stated. This is strictly lower than the welfare loss under traditional Trading if and only

if:

$$\begin{aligned}
& 2\Delta^T W - 2\Delta^{ST} W \geq 0 \\
& \implies \\
& \frac{1}{\beta_1 + \beta_2} \frac{\beta_2^2 \sigma_1^2 + \beta_1^2 \sigma_2^2 + 2\beta_1 \beta_2 \rho \sigma_1 \sigma_2}{\gamma \beta_1 + \gamma \beta_2 + \beta_1 \beta_2} - \frac{\beta_1 + \beta_2}{\gamma \beta_1 + \gamma \beta_2 + \beta_1 \beta_2} \frac{(1 - \rho^2) \sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2 - 2\rho \sigma_1 \sigma_2} \geq 0 \\
& \implies \\
& (\sigma_1^2 + \sigma_2^2 - 2\rho \sigma_1 \sigma_2)(\beta_2^2 \sigma_1^2 + \beta_1^2 \sigma_2^2 + 2\beta_1 \beta_2 \rho \sigma_1 \sigma_2) - (1 - \rho^2)(\beta_1^2 + \beta_2^2 + 2\beta_1 \beta_2) \sigma_1^2 \sigma_2^2 \geq 0 \\
& \implies \\
& [(\beta_2 \sigma_1^2 - \beta_1 \sigma_2^2) + (\beta_1 - \beta_2) \rho \sigma_1 \sigma_2]^2 \geq 0,
\end{aligned}$$

which is always true.

Q.E.D.

PROOF OF PROPOSITION 3:

Proof. We derived quantity derivations under both policies. Prices are equal in both regions, so without loss of generality we can solve for price deviations in region 1:

$$\begin{aligned}
\Delta^T p_1 &= \frac{\beta_2 \theta_1 + \beta_1 \theta_2}{\beta_1 + \beta_2} \\
\Delta^{ST} p_1 &= \frac{\delta \beta_2 \theta_1 + \beta_1 \theta_2}{\beta_1 + \delta \beta_2}.
\end{aligned}$$

Thus:

$$\begin{aligned}
\mathbb{E} [(\Delta^T p)^2] &= \frac{\beta_2^2 \sigma_1^2 + \beta_1^2 \sigma_2^2 + 2\beta_1 \beta_2 \rho \sigma_1 \sigma_2}{\beta_1^2 + \beta_2^2 + 2\beta_1 \beta_2} \\
\mathbb{E} [(\Delta^{ST} p)^2] &= \frac{\delta^2 \beta_2^2 \sigma_1^2 + \beta_1^2 \sigma_2^2 + 2\delta \beta_1 \beta_2 \rho \sigma_1 \sigma_2}{\beta_1^2 + \delta^2 \beta_2^2 + 2\delta \beta_1 \beta_2}.
\end{aligned}$$

Writing these out, we obtain:

$$\mathbb{E} [(\Delta^{ST} p)^2] < \mathbb{E} [(\Delta^T p)^2] \iff (\delta - 1) [\beta_2 (\sigma_1^2 - \rho \sigma_1 \sigma_2) - \beta_1 (\sigma_2^2 - \rho \sigma_1 \sigma_2)] < 0.$$

We now invoke Proposition 2 and establish that this condition is always satisfied. Q.E.D.